DOCUMENT RESUME

ED 101 239 CG 009 474

AUTHOR Remer, Rory; Sease, William
TITLE The Development of a Criterion Instrument for Counselor Selection.
PUB DATE [74]
NOTE 20p.; Paper presented at the Annual Meeting of the American Psychological Association (82nd, New Orleans, Louisiana, August 1974)

EDRS PRICE MF-$0.76 HC-$1.58 PLUS POSTAGE
DESCRIPTORS Counseling Effectiveness; *Counselor Performance; Decision Making; *Interpersonal Competence; *Performance Criteria; *Personnel Selection; Research Projects; Speeches; *Test Validity

ABSTRACT

A measure of potential performance as a counselor is needed as an adjunct to the information presently employed in selection decisions. This article deals with one possible method of development of such a potential perf rmance criterion and the steps taken, to date, in the attempt to val ate it. It includes: the overall effectiveness of the instrument relative to the discrimination of counselor interpersonal competence; the possible modes of presentation of the standardized problem-situations, typescript, audio and audiovisual; the relative cost and the relative effectiveness of each mode of presentation; and results of the convergent-discriminant validation attempt. (Author)

THE DEVELOPMENT OF A CRITERION INSTRUMENT FOR COUNSELOR SELECTION

Rory Remer
Assistant Professor
Educational Foundations
University of Wyoming

William Sease
Associate Professor
Counselor Education
University of Colorado, Denver

Introduction

The purpose of this research effort was to develop an instrument which could be used as a predictive measure of counseling competence based upon performance criteria.

It is generally conceded that the standard academic criteria commonly employed for graduate school admission serve as poor predictors for actual counseling performance. Such measures are useful predictors with respect to other aspects of the counselor's role and functions, e.g., the ability to communicate intelligently, the ability to conduct research, etc., but do not indicate interpersonal competence. Efforts to construct a profile of the successful counselor from the usual array of paper and pencil inventories--interests, personality, values, biographical data, etc.,--have been little more rewarding (Patterson, 1967 and Whiteley 1968, 1969).

An empirically derived instrument which can provide a direct measure of counselor competence in realistic performance-type situations within the limits of reasonable cost and time expenditure is needed. Consequently, the rationale for the present study was based upon a work sample approach utilizing simulated problem-solving situations. Materials were developed, tested and compared across three modes of presentation: typescript, audio and audiovisual.

In the absence of any established external criteria for evaluating counselor competence the rationale depends heavily upon construct validity. The present article focuses especially upon the procedural aspects of instrument development.

Since preliminary results are sufficiently encouraging subsequent cross and longitudinal validation in this and other settings are needed. With the further development of evidence of concurrent and predictive validity, the instrument, or a modified version of it, may serve as an indicator of interpersonal potential in a counseling-type situation.

PROCEDURES

Instrument Development

When potential performance on a complex task must be assessed some sample of a subject's present performance on that task can be the best indicator of his potential (Cronbach, 1970). According to Boyd and Shimberg (1971) when such a sampling can lead to injurious consequences for those involved, it is best to simulate the task. They indicate also that the more precisely the behavior to be assessed can be specified, the better the probability of obtaining a valid measurement. The assessment of potential counselor competence, therefore, calls for a simulated work sample approach.

An instrument to measure the interpersonal competencies needed by a counselor was developed employing the simulated work sample concept. The formulation of the competencies desired of a counselor into specific, measurable behavioral terms; the simulation of a counseling interview to approximate as closely and safely as possible, the actual situation to be encountered in the field; and the methods by which standardized stimuli could be presented and reliable responses obtained in an inexpensive way were given particular attention in the course of instrument development.

Behaviorally defined competencies outlined in the "Ability to Relate to Clients" section of the Application to the California Commission on Teacher Preparation and Licensing for Approval of Stanford University's School of Education Counseling Psychology Program (1972, p. 7) were employed in a slightly modified form. The modified comptencies were:

1) Relate to the client in his own language (understanding)
2) Establish trust and confidence:
    a) Post Hoc Judgment
    b) Does not discuss confidences
3) Have rapport with client:
    a) Accurate Empathy
    b) Personal Communication
4) Look at client as an individual (do not prejudge client)
5) Build up client's confidence (in his ability to deal with his own problems)

While good counseling requires more than just these competencies, these seven dimensions tap the interpersonal aspects required. Since no adequate indicator of interpersonal potential exists which can be used by counselor educators in screening applicants, the instrument was designed to produce a measure of each subject's interpersonal competence as defined by the above modified competencies.

In order to simulate a counseling interview as closely as possible, actual client problem presentations were solicited from counselors. The thirty-five problem presentations, five for each competence, were chosen to be representative of the type encountered when working with clients ranging from high school age through adulthood.

Videotaped interviews were made with actors portraying clients presenting the problems to a counselor. These videotape presentations simulated counseling situations without the danger inherent in actual interaction. They also provided a standard stimulus to which each applicant could react. Presentations in typescript form were also prepared. Included were both problem presentations and a short non-affective, behavioral description of the client.

The cost of preparation and ease of administration of the two different modes of presentation were noted. The typescript was less expensive and easier to administer. Two other modes of presentation, audiotape and timed typescript were also prepared for the purpose of determining comparative discriminating quality.

A multiple choice response form was prepared for use with all four modes of problem presentation. There were thirty-five sets of responses, five corresponding to each of the interpersonal competencies. The choices were of four types: a response designed to exemplify the competence desired to a high degree; a response which was intended to exemplify the competence to a lesser degree; a third response diametrically opposed to the intent of the competence; and a response which was intended to address itself only to the conversational aspects of the problem. The responses within each set were then randomly ordered to prevent any inadvertent response order biases.

The typescript problem presentations and corresponding multiple choice response sets which comprised the instrument were then administered to five behavioral counseling experts from across the country in an attempt to validate the response sets. The correctness of response

was assumed to hold across the different modes of problem presentation. Each expert allotted a total of ten points among the four responses in each set in accord with his judgment as to the appropriateness of the responses to the counseling situation described. Each could allot the total ten points to any alternative or any amount to each of the four responses as long as only ten points were allotted per problem.

One of two criteria was employed to judge correctness of response. The first criteria was a clear majority, 60 percent, of the points allotted by the judges. If the range of judge agreement was between fifty and sixty percent and the question was of high difficulty the response also was designated correct. In the case of a less distinct decision the response or responses were rewritten and resubmitted for judgment. Twenty-nine of the questions had response sets which reached criterion on the first attempt. The remaining six attained criterion on the first iteration. Responses were analyzed to determine interrater reliability using the Hoyt method. A value of .93 was obtained.

No attempt was made to induce a time-pressure situation in any mode of administration. Administrators of the audiotape and videotape modes were instructed to allow sufficient time for all subjects to answer before presenting the next stimulus situation. The timed typescript subjects were allowed a total of seventy minutes for responding to the thirty-five written situations. Subjects experienced no difficulty with the time constraints in any group. The response time required in the audio and video groups was from 25-45 seconds per question. Most subjects in the timed typescript group completed the task within sixty minutes.

Construct Validation

The difficulty in validating an instrument designed to measure a construct is that no immediate external criterion exists with which a convergent validity index may be computed. An indication is needed to determine whether the instrument is functioning well enough to warrant further investigation and/or to justify the use of longitudinal validation procedures. In order to obtain information on which to base such a decision, the "Known Groups" method of construct validation (Scott, 1968, p. 253) was employed. The "Known Groups" procedure while not providing conclusive evidence does give an indication of the instrument's worth. The procedure consists of administering the instrument to a number of groups of subjects known to differ along a continuum with respect to the trait the instrument purports to measure. If the instrument does in fact distinguish among the groups in the predicted manner there is evidence of construct validity.

Correlations between total test score and the admissions criteria presently employed--years of teaching experience, years of counseling experience, undergraduate GPA, graduate GPA, hours of counseling coursework, GRE-Quantitative score, GRE-Verbal score, age, sex and marital status--were also produced. These were used to form a partial convergent-discriminant validity matrix (Cronbach and Meehl, 1955, and Campbell and Fiske, 1959). This information added to the knowledge both of the instrument and of the present criteria used for admissions.

Subjects

Four groups of subjects varying in their ability to relate to others were chosen. The first two groups composed of twenty subjects

each were comprised of professional counselors with at least two years work in the field and practicum students with at least one semester of practicum experience. These subjects represented the high end of the ability continuum. The middle was represented by a group of twenty-one students from a Foundations of Guidance Course at the University of Colorado. This group was similar in composition to applicants to counselor training programs. Twenty-one graduate physics students constituted the low ability level group because of their self-selection into a non-person oriented profession.

The 82 subjects were randomly assigned to mode of presentation within ability level grouping. The audiovisual and the untimed typescript groups consisted of twenty-one subjects each while the other two modality groups consisted of twenty subjects each. The instrument was then administered, scored, and the results analyzed using a two-way Analysis of Variance and Schéffé Multiple Comparison Procedures.

Cost-Benefit Analysis

In order to obtain information concerning the differential effectiveness of the various modes of stimulus presentation, tables were produced based upon the two possible types of incorrect selection decisions. For a number of different cut-off points and several definitions of those acceptable for counselor training percentages of hits, of misses and of false positives were computed for each mode of presentation and for all modes combined.

Rough ratios of cost of presentation of the various modes, including preparation, administration, equipment necessary, etc. were calculated. This information combined with the effectiveness tables

allow an administrator of a counseling program to take into account the importance assigned to each type of selection error with regard to his own program as well as the program's size, costs in personnel and costs of the various modes of presentation. Discriminating quality of each mode of presentation can then be balanced against its relative costs.

RESULTS

Construct Validity

The results of the analysis of variance (ANOVA) of the total test scores of all subjects indicated that the instrument possessed significant ($p < .001$) discriminating qualities although the various modes of presentation did not differ in quality of discrimination.

Insert Table I here

Further analysis of the mean differences employing the Schéffé Multiple Comparisons technique showed that the instrument discriminated among the groups as predicted.

Insert Table II here

Insert Table III here

The graph of the relative distribution of the total test scores for each of the ability level groups appears in Figure I.

Insert Figure I here

The only significant correlations found between the supplementary data collected on all subjects--years of teaching experience, GPA, GRE scores, etc.--and total test score were those calculated for years of counseling experience (.28), for hours of counseling course work (.56), GRE-Quantitative Score (-.41) and for sex (.36). These correlations could be inflated, particularly the sex-total test score correlation, due to composition differences between the high and low ability level groups.

Insert Table IV here

These findings are consistent with those found by Rank (1966) in a study of the relationship between counselor competence and counselor perception which employed stimulus presentations similar in type to those developed for use in the present study.

Cost-Benefit Analysis

A rough cost calculation provided a ratio of 4:2:1. Videotape mode proved to be twice as costly as audiotape mode and four times as expensive as typescript mode. In view of the lack of a significant difference in discrimination by various modes of presentation, the differential cost incurred by using audiotape or videotape presentation does not seem warranted.

CONCLUSIONS AND RECOMMENDATIONS

1. The results support the contention that the development of a standardized instrument for counselor selection, based upon simulated problem situations, is both practical and feasible.

2. The discriminating efficiency regarding interpersonal competence shown by the instrument, i.e., its ability to discriminate among different groups with respect to their ability to relate to others, is sufficiently encouraging to warrant its further consideration and longitudinal validation. While it is true that the evidence provided is not sufficient to indicate conclusive construct validity, the instrument does meet the requirements necessary for further investigation into its validity.

3. Contrary to popular expectations which have been generated by the promise of more sophisticated technology--videotape simulation--the present findings fail to confirm the superiority of this medium over simple typescript situations at a level sufficient to justify its use in this type of testing situation. The authors have observed from the literature that there is growing uncertainty regarding the relative advantages and disadvantages of the audiovisual medium, possibly due to the compounding effects of the multiplicity and types of cues to which the subject is exposed. This is a problem, however, which goes beyond the scope of the present study. See Dwyer (1970), Fast (1971) and Birdwhistell (1970) for further elaboration.

4. No measures of counselor competence exist which have been proved both reliable and valid. Development is needed in the competence area before either concurrent or predictive validity can be assessed accurately.

5. While it by no means measures all competencies required of a counselor, the results indicate that the instrument discussed can add a dimension to the selection procedures not at present considered, i.e., the potential of the candidate to develop the interpersonal skills needed by a counselor.

NOTE: The Potential Interpersonal Competence Scale (PICS) has been copyrighted and an attempt is being made to keep it secure. For further information concerning the instrument itself or any of the tables mentioned but not presented, please feel free to contact the authors.

REFERENCES

Application to the California Commission on Teacher Preparation and Licensing for the Approval of Stanford University's School of Education Counseling Psychology Program, 1972, Draft Proposal.

Birdwhistell, R. L., Kinesics and Context. Philadelphia: University of Pennsylvania Press, 1970.

Boyd, J. L. Jr., and Shimberg, B., "Handbook of Performance Testing: A Practical Guide for Test Makers," ERIC, 1971, EDØ 52220.

Campbell, D. T., and Fiske, D. W., "Convergent and Discriminant Validation by Multitrait-Multimethod Matrix," Psychological Bulletin, 1959, 56:81-105.

Cronbach, L. J., Essentials of Psychological Testing. New York: Harper and Row, 1970.

Cronbach, L. J., and Meehl, P. E., "Construct Validation in Psychological Tests," Psychological Bulletin, 1955, 52:281-302.

Dwyer, F. M. Jr., "Exploration Studies in Effectiveness of Visual Illustrations," Audiovisual Communication Review, 1970, 18:235-249.

Fast, J., Body Language. New York: Pocket Books, 1971.

Patterson, C. H., "The Selection of Counselors," in Research in Counseling, J. M. Whiteley (ed.), Columbus, Ohio: Charles E. Merrill, 1967, pp. 69-102.

Rank, R. C., "Counseling Competence and Perceptions," Personnel and Guidance Journal, December, 1966, p. 359.

Scott, W. A., "Attitude Measurement," The Handbook of Social Psychology, Lindzey, G., and Aronson, E. (eds.), Volume 2, Reading, Massachusetts: Addison-Wesley Publishing Company, 1968, pp. 204-273.

Whiteley, J. M., "Counselor Education: A Critical Review of the Literature, 1965-1968," ERIC, 1968, ED0 26674.

Whiteley, J. M., "Counselor Education," Review of Educational Research, 1969, 2:173-187.

TABLE I

TWO-FACTOR 4x4 ANALYSIS OF VARIANCE
MODE OF PRESENTATION BY
ABILITY LEVEL GROUPING

Source of Variation	SS	df	MS	F
Mode	579.62	3	193.21	1.37
Level	11631.10	3	3877.03	27.55*
Mode X Level	1194.65	9	132.74	.94
Residual	9288.53	66	140.74	

* Significant $p < .001$

TABLE II

TABLE OF MEAN DIFFERENCES

Group	Expert	Practicum	Applicant	Physics
Means	115.70 (9.9)[a]	117.10 (11.5)[a]	101.05 (13.3)[a]	87.33 (12.7)[a]
Expert	0	-1.4	14.65*	28.32*
Practicum		0	16.05*	29.72*
Applicant			0	13.67*

* Significant $p < .001$

a-standard deviations in parentheses

TABLE III

SCHEFFÉ TESTS

Scheffé Comparisons

Comparison	$C_{\mu E}$	$C_{\mu PR}$	$C_{\mu AP}$	$C_{\mu PH}$	$\psi^2/\hat{\sigma}^2_{\psi(J-1)}$	F
$\psi 1$	1	-1	0	0	.14	F(1,38)
$\psi 2$	0	0	1	-1	13.94	F(1,40)*
$\psi 3$	1	0	-1	0	15.62	F(1,39)*
$\psi 4$	1	0	0	-1	64.29	F(1,39)*
$\psi 5$	0	1	0	-1	58.38	F(1,39)*
$\psi 6$	0	1	-1	0	18.75	F(1,39)*
$\psi 7$	½	½	-1	0	11.62	F(2,58)*
$\psi 8$	½	½	0	-1	41.54	F(2,58)*
$\psi 9$	½	½	-½	-½	23.88	F(3,78)*

MS_W = 140.73535

* Significant $p < .001$

TABLE IV

CORRELATIONS WITH TOTAL TEST SCORE

Variable	Correlation	Mean	S.D.
Level Value	-.67**	2.52	1.12
Quartile	.94**	2.50	1.11
Years of Teaching Experience	.13	3.18	3.49
Years of Counseling Experience	.28*	2.15	3.54
Undergraduate GPA	-.10	3.08	.40
Graduate GPA	.21	3.65	.25
Hours of Counseling Course Work	.56**	15.70	19.90
GRE-Quantitative	-.41**	595.54	134.02
GRE Verbal	.05	591.69	114.47
Age	.20	30.06	6.05
Sex (Female-1, Male-0)	.36**	.38	.49
Marital Status (Married=2, Single=1)	.14	1.59	.84

* Significant $p < .05$

** Significant $p < .001$

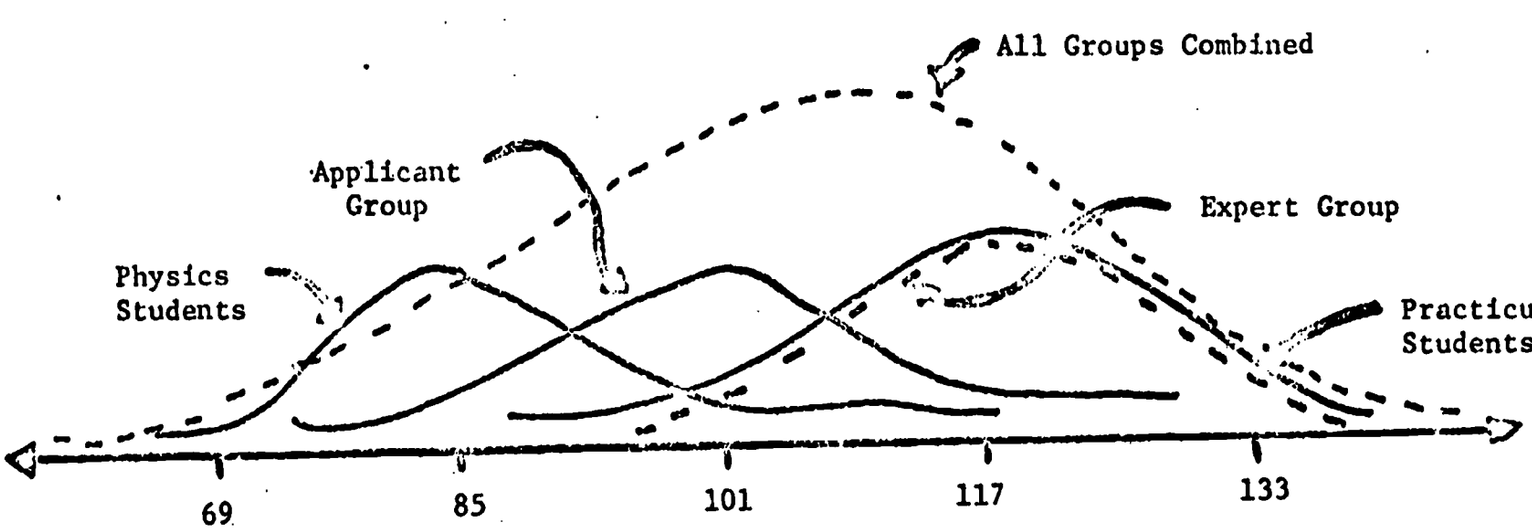

Figure I

Distribution of Test Scores by Group